

# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 2002-245066

(43)Date of publication of application : 30.08.2002

(51)Int.Cl. G06F 17/30  
G10K 15/02  
G10L 15/00  
G10L 15/28  
H04N 5/76

(21)Application number : 2001-041517

(71)Applicant : SONY CORP

(22)Date of filing : 19.02.2001

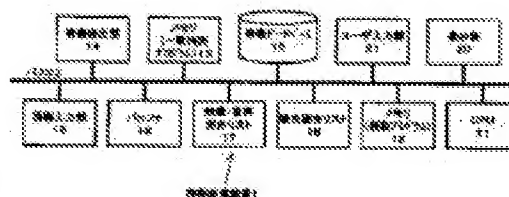
(72)Inventor : HIRAI JUN

## (54) DEVICE AND METHOD FOR RETRIEVING INFORMATION AND STORAGE MEDIUM

### (57)Abstract:

**PROBLEM TO BE SOLVED:** To retrieve contents by utilizing the featured values of the both video and voice in AV(audio visual) contents.

**SOLUTION:** A database where voice featured values are connected to a voice identifying code, a database where video featured values are connected to a video identifying code and an AV contents table where the voice identifying code is connected to the video identifying code are prepared so as to retrieve the video and the voice in parallel. When the video is retrieved, the AV contents table is referred to concerning the video identifying code so as to pickup the corresponding voice identifying code. The code is given to the voice retrieval and, then, the retrieval is completed when collation is obtained.



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号  
特開2002-245066  
(P2002-245066A)

(43) 公開日 平成14年8月30日 (2002.8.30)

(51) Int.Cl. <sup>7</sup>	識別記号	F I	テーマコード* (参考)	
G 0 6 F 17/30	2 1 0 1 7 0	G 0 6 F 17/30	2 1 0 C	5 B 0 7 j 1 7 0 D 1 7 0 E
G 1 0 K 15/02		G 1 0 K 15/02		5 C 0 5 2 5 D 0 1 5
G 1 0 L 15/00		H 0 4 N 5/76	B	
審査請求 未請求 請求項の数 9 O L (全 13 頁) 最終頁に続く				

(21) 出願番号 特願2001-41517(P2001-41517)

(22) 出願日 平成13年2月19日 (2001.2.19)

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 平井 純

東京都品川区北品川6丁目7番35号 ソニー株式会社内

(74) 代理人 100101801

弁理士 山田 英治 (外2名)

Fターム(参考) 5B075 ND04 ND12 ND14 ND35 NK02

NK43 NK54 PP12 PP22 QM01

UU35 UU36

5C052 AA01 AB10 AC08 DD10

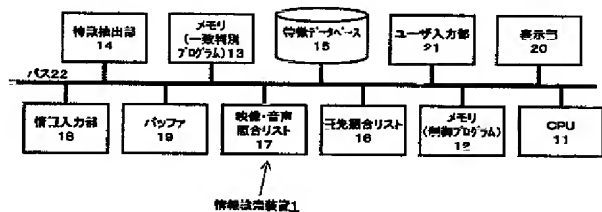
5D015 KK03

(54) 【発明の名称】 情報検索装置及び方法、並びに記憶媒体

(57) 【要約】

【課題】 AVコンテンツの映像と音声の両方の特徴量を利用してコンテンツ検索を行う。

【解決手段】 音声の特徴量と音声識別コードを連結したデータベースと、映像の特徴量と映像識別コードを連結したデータベースと、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルを用意する。映像の検索と音声の検索を並行して実行する。例えば、映像の検索で見つかった場合にその映像識別コードをAVコンテンツ・テーブルで参照して、該当する音声識別コードを引き出す。このコードを音声の検索に与えて、照合がとれれば検索は完了する。



## 【特許請求の範囲】

【請求項1】映像データと音声データの組み合わせからなるAVコンテンツを検索する情報検索装置であって、音声の特徴量と音声識別コードを連結した音声データベースと、映像の特徴量と映像識別コードを連結した映像データベースと、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルと、音声データベースを用いて音声の検索を行う音声データ検索手段と、映像データベースを用いて映像の検索を行う映像データ検索手段と、前記音声データ検索手段及び前記映像データ検索手段により抽出された音声識別コード及び映像識別コードに基づいて、前記AVコンテンツ・テーブルから該当するAVコンテンツを決定するAVコンテンツ決定手段と、を具備することを特徴とする情報検索装置。

【請求項2】前記音声データ検索手段及び前記映像データ検索手段は並行して検索を実行し、前記映像データ検索手段で先に見つかった場合には、前記AVコンテンツ決定手段は、その映像識別コードを前記AVコンテンツ・テーブルで参照して該当する音声識別コードを引き出し、該音声識別コードを音声データ検索手段に与えて照合をとり、前記音声データ検索手段で先に見つかった場合には、前記AVコンテンツ決定手段は、その音声識別コードを前記AVコンテンツ・テーブルで参照して該当する映像識別コードを引き出し、該映像識別コードを映像データ検索手段に与えて照合をとる、ことを特徴とする請求項1に記載の情報検索装置。

【請求項3】検索対象となるAVコンテンツのコンテンツ識別コード取得手段をさらに備え、前記音声データ検索手段及び／又は映像データ検索手段は、該コンテンツ識別コードに対応するAVコンテンツに対して優先的に照合を行う、ことを特徴とする請求項1に記載の情報検索装置。

【請求項4】前記音声データ検索手段及び／又は映像データ検索手段は、明かに相違するAVコンテンツに対する照合処理を省略する、ことを特徴とする請求項1に記載の情報検索装置。

【請求項5】映像データと音声データの組み合わせからなるAVコンテンツを検索する情報検索方法であって、音声の特徴量と音声識別コードを連結した音声データベースを用いて音声の検索を行う音声データ検索ステップと、映像の特徴量と映像識別コードを連結した映像データベースを用いて映像の検索を行う映像データ検索ステップと、前記音声データ検索ステップ及び前記映像データ検索ス

テップにより抽出された音声識別コード及び映像識別コードに基づいて、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルから該当するAVコンテンツを決定するAVコンテンツ決定ステップと、を具備することを特徴とする情報検索方法。

【請求項6】前記音声データ検索ステップ及び前記映像データ検索ステップは並行して検索を実行し、前記映像データ検索ステップにより先に見つかった場合には、前記AVコンテンツ決定ステップでは、その映像識別コードを前記AVコンテンツ・テーブルで参照して該当する音声識別コードを引き出し、該音声識別コードを音声データ検索ステップに投入して照合をとり、前記音声データ検索ステップにより先に見つかった場合には、前記AVコンテンツ決定ステップでは、その音声識別コードを前記AVコンテンツ・テーブルで参照して該当する映像識別コードを引き出し、該映像識別コードを映像データ検索ステップに投入して照合をとる、ことを特徴とする請求項5に記載の情報検索方法。

【請求項7】検索対象となるAVコンテンツのコンテンツ識別コード取得ステップをさらに備え、前記音声データ検索手段及び／又は映像データ検索ステップでは、該コンテンツ識別コードに対応するAVコンテンツに対して優先的に照合を行う、ことを特徴とする請求項5に記載の情報検索方法。

【請求項8】前記音声データ検索ステップ及び／又は映像データ検索ステップは、明かに相違するAVコンテンツに対する照合処理を省略する、ことを特徴とする請求項5に記載の情報検索方法。

【請求項9】映像データと音声データの組み合わせからなるAVコンテンツを検索する情報検索処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、音声の特徴量と音声識別コードを連結した音声データベースを用いて音声の検索を行う音声データ検索ステップと、映像の特徴量と映像識別コードを連結した映像データベースを用いて映像の検索を行う映像データ検索ステップと、前記音声データ検索ステップ及び前記映像データ検索ステップにより抽出された音声識別コード及び映像識別コードに基づいて、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルから該当するAVコンテンツを決定するAVコンテンツ決定ステップと、を具備することを特徴とする記憶媒体。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、情報検索技術に係り、特に、映像との音楽の各データからなるAVコンテ

ンツを検索する情報検索技術に関する。

【0002】更に詳しくは、本発明は、元データが持つ特徴量に基づくAVコンテンツのための情報検索技術に係り、特に、映像と音楽の両方の特徴量を利用してAVコンテンツを検索する情報検索技術に関する。

【0003】

【従来の技術】昨今、コンピュータ上での演算処理をベースとした情報処理技術が飛躍的に向上している。これに伴い、単なるコンピュータ・データ以外に、映像や音声などの各種のデータ・コンテンツもデジタル化され、コンピュータ・ファイルとして取り扱われるようになってきている。したがって、情報資源の有効活用という観点からも、各種のデータ・コンテンツに関する検索技術は重要度を増している。

【0004】あるいは、コンテンツの配信や放送の分野において、コンテンツ使用量に関して従量課金制を導入した場合や、著作権保護の観点からコンテンツの使用を取り締まる場合など、各コンテンツに対して配信又は放送された回数を計数したり監視しなければならない。例えば、いま配信されたコンテンツが何であったかを特定するために、情報検索技術が適用される。

【0005】例えば、テキスト形式の比較的小サイズのデータ検索であれば、コンピュータの演算能力に頼ってテキストの一致を基調とする全数検索を行うこともできる。これに対し、映像や音楽データからなるAVコンテンツの場合、一般にデータ量が膨大であるため、コンテンツの初めから終わりまで全数検索を行うことは技術的に不可能である。

【0006】音楽情報検索として、楽曲データを採譜して検索対象となる楽曲データとの照合（全数検索）を行うものがある。この方法では、楽曲の初めから終わりまでパターン・マッチングを行う。しかしながら、インデックスを使用しない全数検索は、検索速度が遅いという欠点がある。

【0007】例えば音楽放送においては、音譜ベースでの比較照合ではなく、信号の特徴抽出による監視が既に実施されている。すなわち、ISRC（International Standard Recording Number）のような楽曲のユニーク・コードと対応させた楽曲の特徴量をあらかじめデータベースに登録しておく。ISRCや楽曲の採取には、CD（Compact Disc）などの既存の記録メディアを利用する。そして、放送を受信してその特徴量とデータベースに登録済みの特徴量とを比較照合して、受信した楽曲のISRCを特定する（図6を参照のこと）。

【0008】音楽放送で実施されている同一検索を、AVコンテンツにまでそのまま適用範囲を拡張する場合、音声で行うと無音部分では抽出できないという問題がある。例えば、音楽のみのコンテンツに比べ、映画などでは無音部分が多いので、音楽データの特徴抽出に基づく検索には限界がある。

【0009】一方、映像だけでAVコンテンツを検索することを考えた場合、登場人物がしゃべっているシーンなど映像の内容の変化が少ない部分が多いので、映像データの特徴抽出に基づく検索には限界がある。

【0010】また、映画業界においては、同じ映画でも販売先となるそれぞれの国や地域の公用語などに応じた言語対応、すなわち音声の吹き替え処理が行われる。吹き替え処理の結果、音声ベースでのコンテンツ検索を行うと、同じ映画であっても一致しなくなる。

【0011】

【発明が解決しようとする課題】本発明の目的は、映像との音楽の各データからなるAVコンテンツを高速且つ確実に検索することができる、優れた情報検索装置及び方法を提供することにある。

【0012】

【課題を解決するための手段及び作用】本発明は、上記課題を参酌してなされたものであり、その第1の側面は、映像データと音声データの組み合わせからなるAVコンテンツを検索する情報検索装置又は方法であって、音声の特徴量と音声識別コードを連結した音声データベースと、映像の特徴量と映像識別コードを連結した映像データベースと、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルと、音声データベースを用いて音声の検索を行う音声データ検索手段又はステップと、映像データベースを用いて映像の検索を行う映像データ検索手段又はステップと、前記音声データ検索手段又はステップ、及び、前記映像データ検索手段又はステップにより抽出された音声識別コード及び映像識別コードに基づいて、前記AVコンテンツ・テーブルから該当するAVコンテンツを決定するAVコンテンツ決定手段又はステップと、を具備することを特徴とする情報検索装置又は方法である。

【0013】本発明に係る情報検索装置又は方法は、音声の特徴量と音声識別コードを連結したデータベースと、映像の特徴量と映像識別コードを連結したデータベースと、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルを用意する。そして、映像の検索と音声の検索を並行して実行する。

【0014】例えば、映像の検索で見つかった場合にその映像識別コードをAVコンテンツ・テーブルで参照して、該当する音声識別コードを引き出す。このコードを音声の検索に与えて、照合がとれれば検索は完了する。同様に、音声の検索で見つかった場合にその音声識別コードをAVコンテンツ・テーブルで参照して、該当する映像識別コードを引き出す。このコードを映像の検索に与えて、照合がとれれば検索は完了する。

【0015】放送系などでこのような検索を行う場合、マルチパスの影響や、送信・受信回路の歪や圧縮による歪など、系としてある程度の歪を許容しなければならない場合がある。この場合、完全な同一検索では検索

不能になるので、特徴量の照合時に、ある程度の判定の許容度を与える必要がある。逆に許容度を与えることによって、ある1つの検索対象に対して複数の検索結果が出現する可能性がある。

【0016】検索の高速化を狙った場合、複数の検索結果が出てしまう可能性は必然的に高まる。このような事態を解消するためには、映像と音声の両方を検索対象にして、信頼性を上げるようにすればよい。

【0017】映像で検索して先に見つかった場合には、この映像識別コードをAVコンテンツ・テーブルを参照して、音声の検索に音声識別コードを与えて照合を図る。また、その逆に、音声で検索して先に見つかった場合に、この音声識別コードをAVコンテンツ・テーブルを参照して、映像の検索に映像識別コードを与えて照合を図る。そして、映像及び音声の両方の結果が一致することを見て判断する。映像、音声の検索結果がそれぞれ複数出現する場合には、すべての組み合わせについて判断し、誤差が少ない結果を選択するようにする。

【0018】本発明の第1の側面に係る情報検索装置又は方法は、さらに検索対象となるAVコンテンツのコンテンツ識別コード取得手段又はステップを備えて、前記音声データ検索手段又はステップ、及び／又は、映像データ検索手段又はステップは、該コンテンツ識別コードに対応するAVコンテンツに対して優先的に照合を行うようにしてもよい。

【0019】コンテンツを識別する識別コードなどの情報を入手するメカニズムが用意されている場合であって、しかも改竄のおそれがなく入手できる場合には、信号の照合も必要がない。しかし、コンテンツの著作権者は、再生機や記録機をユーザが改竄したりして偽りの識別コードが生成されることを恐れる。そこで、改竄を確認するための信号の照合を行う必要がある。

【0020】多くの場合は改竄されていないので、コンテンツ識別コードに対応するコンテンツを優先して照合処理することによって、システム全体の情報検索速度を高速化することができる。

【0021】また、前記音声データ検索手段又はステップ、及び／又は、映像データ検索手段又はステップは、明かに相違するAVコンテンツに対する照合処理を省略するようにしてもよい。

【0022】明かに相違することが容易に分かるAVコンテンツに関しては、音声データ検索並びに映像データ検索を行う必要がなく、かかる工程を省略することにより計算機負荷を軽減することができるとともに、より類似するコンテンツに関して誤差を小さくして高精度な照合処理を行うことができる。また、明かに相違するコンテンツに対する余分な照合処理を省くので、システム全体の情報検索速度を高速化することができる。

【0023】本発明の第2の側面は、映像データと音声データの組み合わせからなるAVコンテンツを検索する

情報検索処理をコンピュータ・システム上で実行するように記述されたコンピュータ・ソフトウェアをコンピュータ可読形式で物理的に格納した記憶媒体であって、前記コンピュータ・ソフトウェアは、音声の特徴量と音声識別コードを連結した音声データベースを用いて音声の検索を行う音声データ検索ステップと、映像の特徴量と映像識別コードを連結した映像データベースを用いて映像の検索を行う映像データ検索ステップと、前記音声データ検索ステップ及び前記映像データ検索ステップにより抽出された音声識別コード及び映像識別コードに基づいて、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルから該当するAVコンテンツを決定するAVコンテンツ決定ステップと、を具備することを特徴とする記憶媒体である。

【0024】本発明の第2の側面に係る記憶媒体は、例えば、様々なプログラムコードを実行可能な汎用コンピュータシステムに対して、コンピュータ・ソフトウェアをコンピュータ可読な形式で提供する媒体である。このような媒体は、例えば、CD (Compact Disc) やFD (Floppy Disk)、MO (Magnet-Optical disc) などの着脱自在で可搬性の記憶媒体である。あるいは、ネットワーク（ネットワークは無線、有線の区別を問わない）などの伝送媒体などを經由してコンピュータ・ソフトウェアを特定のコンピュータ・システムに提供することも技術的に可能である。

【0025】このような記憶媒体は、コンピュータ・システム上で所定のコンピュータ・ソフトウェアの機能を実現するための、コンピュータ・ソフトウェアと記憶媒体との構造上又は機能上の協働的関係を定義したものである。換言すれば、本発明の第2の側面に係る記憶媒体を介して所定のコンピュータ・ソフトウェアをコンピュータ・システムにインストールすることによって、コンピュータ・システム上では協働的作用が発揮され、本発明の第1の側面に係る情報検索装置及び方法と同様の作用効果を得ることができる。

【0026】本発明のさらに他の目的、特徴や利点は、後述する本発明の実施例や添付する図面に基づくより詳細な説明によって明らかになるであろう。

【0027】

【発明の実施の形態】以下、図面を参照しながら本発明の実施例を詳解する。

【0028】図1には、本発明の一実施形態に係る情報検索装置1の構成を模式的に示している。以下、同図を参照しながら、情報検索装置1の各コンポーネントについて説明する。

【0029】CPU (Central Processing Unit) 11は、情報検索装置1全体の動作を統括的に制御する中央コントローラであり、オペレーティング・システム (OS) が提供するプログラム実行環境下で、メモリ12に格納された（又はロードされた）制御プログラムや他の

メモリ13に格納された(又はロードされた)一致判別プログラムを実行するようになっている。

【0030】ここで言う一致判別プログラムには、映像の特徴値による映像データベースの照合処理、音声の特徴値による音声データベースの照合処理、映像並びに音声データベースの照合結果に基づくAVコンテンツ・テーブルの照合処理などが含まれる。また、制御プログラムには、一致判別結果に基づくAVコンテンツの特定処理などのAVコンテンツに関する情報検索処理プログラムが含まれる。

【0031】CPU11は、バス22を介して装置1内の各モジュールと相互接続されている。バス22は、アドレス・バス、データ・バス、コントロール・バスを含んだ共通信号伝送路である。

【0032】映像・音声特徴抽出部14は、AVコンテンツから映像及び音声データへのデマルチプレクス処

理、並びに、映像データの特徴抽出処理、音声データの特徴抽出処理を行う。映像・音声特徴抽出部14によって抽出された映像並びに音声の特徴値は、一致判別プログラムに投入される。

【0033】特徴データベース15は、AVコンテンツに関する照合・情報検索処理に利用されるデータベースで構成される。より具体的には、特徴データベース15は、音声の特徴量と音声識別コードを連結した音声データベースと、映像の特徴量と映像識別コードを連結した映像データベースと、音声識別コードと映像識別コードを連結したAVコンテンツ・テーブルとで構成される。このうち映画に関するAVコンテンツ・テーブルの構成例を以下に示しておく。

【0034】

【表1】

映画名	映像識別コード	音声識別コード	AVコンテンツ識別コード
映画1英語版	1	1	1-1
映画2日本語版	2	2	2-2
映画2英語版	2	3	2-3
映画3英語版	3	4	3-4
映画3日本語版	3	5	3-5
映画4英語版	4	6	4-6
映画4日本語版	4	7	4-7
映画4スペイン語版	4	8	4-7

【0035】上記に示したAVコンテンツ・テーブルでは、各レコードは、映画名と、映像識別コードと、音声識別コードと、AVコンテンツ識別コードの組み合わせで構成される。したがって、映像識別コードが判明するとこれにヒットする音声識別コードを絞り込むことができるし、逆に、音声識別コードが判明するとこれにヒットする映像識別コードを絞り込むことができる。また、映像識別コードと音声識別コードの双方を特定すると、このAVコンテンツ・テーブルを参照することによってAVコンテンツ識別コードすなわちAVコンテンツが一意に定まる。

【0036】優先照合リスト16、並びに映像・音声照合リスト17は、情報検索すなわちAVコンテンツの照合処理を行うときの一部の作業データを一時的に保管するために使用される。

【0037】優先照合リスト16は、優先的に照合処理を行うべきAVコンテンツをリストするための記憶モジュールである。例えば、外部からコンテンツ識別コードが渡されたAVコンテンツに関しては、この識別コードにより照合・検索範囲を限定して処理時間を高速化することができるまで、優先照合リスト16に登録する。

【0038】映像・音声照合リスト17は、各映像データ及び音声データについてのデータベース照合結果を一時的に保管するために使用される。例えば、映像データが一致するが対応する音声データが一致しない場合は、映像データの照合処理を行う必要がないので、映像照合

リストにその旨を記入する。また、逆に音声データが一致するが対応する映像データが一致しない場合は、音声データの照合処理を行う必要がないので、音声照合リストにその旨を記入する。また、音声明らかに違う場合には、音声照合リストにその旨を記入して高精度の照合処理をスキップさせるようにするし、映像が明らかに違う場合には、映像照合リストにその旨を記入して照合処理をスキップさせるようにする。

【0039】情報入力部18は、外部機器から情報を入力する機能モジュールである。ここで入力される情報には、映像データと音声データがマルチプレクスされたAVコンテンツや、そのコンテンツ識別コードなどが含まれる。また、情報を入出力する媒体として、LAN (Local Area Network) やインターネットなどのコンピュータ・ネットワーク、CD (Compact Disc) やDVD (Digital Versatile Disc) などの可搬型記録メディアなどを挙げることができる。情報入力部18を介して外部の装置から供給された情報は、例えば、バッファ・メモリ19上に一時格納される。

【0040】表示部20は、CPU11による演算結果を画面出力してユーザに対して情報の視覚的なフィードバックを行う。また、ユーザ入力部21は、例えばキーボードやマウス(いずれも図示しない)などの入力装置で構成され、ユーザからのコマンドやデータの入力を受容する。

【0041】なお、情報検索装置1は、例えば一般的な

計算機システムを利用して構成することができる。かかる計算機システムの例は、米IBM社のPC/AT (Personal Computer/Advanced Technology) 互換機又はその後継機である。

【0042】図2には、情報検索装置1上で実行されるAVコンテンツの検索処理についての概略的な手順を、フローチャートの形式で示している。以下、このフローチャートを参照しながら、本実施形態に係るAVコンテンツの検索処理の流れを概略的に説明する。

【0043】まず、検索対象となるAVコンテンツをデマルチプレクスして、映像信号と音声信号に分離する(ステップS1)。なお、図示しないが、必要に応じて映像信号と音声信号をそれぞれ圧縮を解凍する。

【0044】次いで、映像信号と音声信号からそれぞれ特徴量を抽出して、特徴データベース15にあらかじめ登録された各データと照合して、コンテンツ識別コードを抽出する。ステップS2Aの音声照合とステップS2Bの映像照合は並行して行う。コンテンツ識別コードは、特徴データベース15中のAVコンテンツ・テーブル(前述)を参照して、音声であればそれに組み合わされる映像の識別コードを抽出し(ステップS3A)、映像であればそれに組み合わされる音声の識別コードを抽出する(ステップS3B)。

【0045】抽出された音声識別コードは音声照合2(ステップS4A)に渡され、そのコンテンツ識別コードに該当する特徴量を音声データベースと照合することで、該当する音声のコンテンツ識別コードを確認あるいは選別する。同様に、抽出された映像識別コードは映像照合2(ステップS4B)に渡され、そのコンテンツ識別コードに該当する特徴量を映像データベースと照合することで、該当する映像のコンテンツ識別コードを確認あるいは選別する。

【0046】音声照合2あるいは映像照合2で得られたコンテンツ識別コードのうち早期に得られたものを採用するように決定論理(ステップS5)を組むことによって、AVコンテンツの識別の高速化を図ることができる。

【0047】[表1]に示したAVコンテンツ・テーブルを使用した場合、例えば、映像照合により映像識別コード(4)を抽出してAVコンテンツ・テーブルを参照すると、対応する音声識別コードとして(6, 7, 8)を取り出すことができる。これら音声識別コードを音声照合2に投入して、そこで照合が合致することによってAVコンテンツ識別コード(例えば、4-7)が判断される。

【0048】伝送系の歪、伝送時のマルチパスや圧縮や記録再生時の歪など、完全な同一検索ではなく、歪に対応するため、ある程度の許容度を照合時に認める場合には、ただ1つの識別コードに対して照合時に2つのコンテンツが対応してしまうことがある。このような場合に

は、音声照合2と映像照合2が一致した識別コードを採用するように、決定論理を構成して、信頼性の向上を図ることができる。

【0049】図2に示したフローチャートにおいて、点線で記したパスを設けて、音声照合2又は映像照合2から識別コードが得られなかった場合には、映像照合又は音声成功空の識別コードを直接利用して、「音声は正常ではないようですが映像から識別コードはXXXXXと読めます」、あるいは「映像が正常ではないようですが音声から識別コードはXXXXXと読めます」などのユーザ・フィードバックを、表示部20などを介して行うようにしてもよい。

【0050】図3には、本実施形態に係る情報検索装置1上で実行されるAVコンテンツの検索処理の具体的な手順の一例を、フローチャートの形式で示している。同図に示す処理手順は、照合に誤差を許容しない場合の例である。

【0051】まず、検索対象となるAVコンテンツをデマルチプレクスして、映像信号と音声信号に分離する(ステップS11)。なお、図示しないが、必要に応じて映像信号と音声信号をそれぞれ圧縮を解凍する。

【0052】次いで、映像信号と音声信号からそれぞれ特徴量を抽出する(ステップS12)。そして、映像並びに音声それぞれの特徴値を特徴データベース15にあらかじめ登録された各データと照合して(ステップS13)、一致するか否かを判断する(ステップS14)。映像及び音声の照合の速度はコンテンツによって区々である。本実施形態では、映像及び音声の照合処理をそれぞれベストエフォートで同時並行的に実行する。

【0053】映像並びに音声のいずれについても一致を見出せなかった場合には(ステップS14)、特徴データベース中の次のコンテンツに進んで(ステップS20)、同様に映像・音声の特徴量の照合を繰り返し実行する。そして、特徴データベース中で一致するコンテンツが見つからなかった場合には、該当コンテンツがない旨の表示を表示部20から出力してユーザに通知し(ステップS21)、本処理ルーチン全体を終了する。

【0054】映像について一致するデータが見つかった場合(ステップS15)、音声照合リストを参照して、この映像データに対応する音声について、一致するものがあつたか否かを判別する(ステップS18)。対応音声について一致するものがない場合には、映像照合リスト中に当該映像は一致しない旨を記録した後(ステップS19)、ステップS13に復帰して、対応映像のコンテンツにジャンプして、映像の特徴データベースと比較照合する。

【0055】また、音声について一致するデータが見つかった場合(ステップS15)、映像照合リストを参照して、この音声データに対応する映像について、一致するものがあつたか否かを判別する(ステップS16)。



対応映像については一致するものがない場合には、音声照合リスト中に当該音声は一致しない旨を記録した後（ステップS17）、ステップS13に復帰して、対応音声のコンテンツにジャンプして、音声の特徴データベースと比較照合する。

【0056】判断ブロックS16又はS18の結果、映像及び音声について一致するAVコンテンツが見つかった場合、そのコンテンツ識別コードを検索結果として出力して、本処理ルーチン全体を終了する。

【0057】また、図4には、本実施形態に係る情報検索装置1上で実行されるAVコンテンツの検索処理の具体的な手順の他の例を、フローチャートの形式で示している。同図に示す処理手順は、何らかの手段により与えられたコンテンツ識別コードを利用して検索し、照合に誤差を許容しない場合の例である。

【0058】まず、何らかの手段により、検索対象となるAVコンテンツのコンテンツ識別コードを取得して（ステップS31）、これを優先照合リスト16に登録する（ステップS32）。本実施形態に係る情報検索装置1は、例えば検索対象としてのAVコンテンツを絶え間なく入力し続けるような場合、優先照合リスト16に登録されたAVコンテンツを優先的に照合処理する。

【0059】次いで、優先照合リスト16に登録された再優先のコンテンツ識別コードに対応するAVコンテンツをデマルチプレクスして、映像信号と音声信号に分離する（ステップS33）。なお、図示しないが、必要に応じて映像信号と音声信号をそれぞれ圧縮を解凍する。

【0060】次いで、映像信号と音声信号からそれぞれ特徴量を抽出する（ステップS34）。そして、映像並びに音声それぞれの特徴値を特徴データベース15にあらかじめ登録された各データと照合して（ステップS35）、一致するか否かを判断する（ステップS36）。映像及び音声の照合の速度はコンテンツによって区々である。本実施形態では、映像及び音声の照合処理をそれぞれベストエフォートで同時並行的に実行する。

【0061】映像並びに音声のいずれについても一致を見出せなかった場合には（ステップS36）、特徴データベース中の次のコンテンツに進んで（ステップS42）、同様に映像・音声の特徴量の照合を繰り返し実行する。そして、特徴データベース中で一致するコンテンツが見つからなかった場合には、該当コンテンツがない旨の表示を表示部20から出力してユーザに通知し（ステップS43）、本処理ルーチン全体を終了する。

【0062】映像について一致するデータが見つかった場合（ステップS37）、音声照合リストを参照して、この映像データに対応する音声について、一致するものがあつたか否かを判別する（ステップS40）。対応音声について一致するものがない場合には、映像照合リスト中に当該映像は一致しない旨を記録した後（ステップS41）、ステップS35に復帰して、対応映像の

コンテンツにジャンプして、映像の特徴データベースと比較照合する。

【0063】また、音声について一致するデータが見つかった場合（ステップS37）、映像照合リストを参照して、この音声データに対応する映像について、一致するものがあつたか否かを判別する（ステップS38）。対応映像については一致するものがない場合には、音声照合リスト中に当該音声は一致しない旨を記録した後（ステップS39）、ステップS35に復帰して、対応音声のコンテンツにジャンプして、音声の特徴データベースと比較照合する。

【0064】判断ブロックS38又はS40の結果、映像及び音声について一致するAVコンテンツが見つかった場合、そのコンテンツ識別コードを検索結果として出力して、本処理ルーチン全体を終了する。

【0065】検索対象となるAVコンテンツのコンテンツIDを改竄のおそれなく入手できる場合、信号の照合も必要ない。しかしながら、コンテンツの著作権者は、再生機や記録機を改造したりして偽りのIDが作られることをおそれる。そこで、その改竄を確認するため、信号の照合を行う必要がある。多くの場合は改竄されていないので、上述したように優先照合リスト16中のコンテンツを優先して照合して、一致していればそれでコンテンツ検索を終了することができるので、検索速度を大幅に高速化することができる。

【0066】また、図5には、本実施形態に係る情報検索装置1上で実行されるAVコンテンツの検索処理の具体的な手順の他の例を、フローチャートの形式で示している。同図に示す処理手順は、特徴データベースに登録されたAVコンテンツのうち検索対象となるAVコンテンツとは明らかに相違するものを除外して照合処理することにより、高速化を図った例である。

【0067】まず、検索対象となるAVコンテンツをデマルチプレクスして、映像信号と音声信号に分離する（ステップS51）。なお、図示しないが、必要に応じて映像信号と音声信号をそれぞれ圧縮を解凍する。

【0068】次いで、映像信号と音声信号からそれぞれ特徴量を抽出する（ステップS52）。そして、映像並びに音声それぞれの特徴値を特徴データベース15にあらかじめ登録された各データと照合して（ステップS53）、検索対象となるAVコンテンツとは明らかに相違するか否かを判断する（ステップS54）。映像及び音声の照合の速度はコンテンツによって区々である。本実施形態では、映像及び音声の照合処理をそれぞれベストエフォートで同時並行的に実行する。ここでは、コンテンツの相違を検証することのみを目的とするので、ステップS53では比較的大きな誤差を認めた照合処理を行えばよい。

【0069】ステップS54において、明らかに相違すると判断された場合、映像又は音声のうちいずれが明ら



かに相違するかをさらに判別する（ステップS55）。

【0070】映像に関して明らかに相違する場合には、映像照合リスト中に当該映像は一致しない旨を記録した後（ステップS56）、ステップS53に復帰して、対応映像の照合処理をスキップする。

【0071】また、音声に関して明らかに相違する場合には、音声照合リスト中に当該音声は一致しない旨を記録した後（ステップS57）、ステップS53に復帰して、対応音声の照合処理をスキップする。

【0072】検索対象となるAVコンテンツと明らかに相違する訳ではない場合には、特徴データベース中の次のコンテンツに進んで（ステップS58）、同様に検索対象となるAVコンテンツと明らかに相違しないか否かを判別する処理を繰り返し実行する。

【0073】特徴データベース中のすべてのデータについて、検索対象となるAVコンテンツとの相違を検証した後、映像並びに音声それぞれの特徴値を特徴データベース15にあらかじめ登録された各データと照合して（ステップS59）、検索対象となるAVコンテンツとは相違するか否かを判断する（ステップS60）。映像及び音声の照合の速度はコンテンツによって区々である。本実施形態では、映像及び音声の照合処理をそれぞれベストエフォートで同時並行的に実行する。ここでは、データの絞込みを目的とするので、ステップS59では先行ステップS53よりも誤差を小さくした照合処理を行う。

【0074】ステップS60において、相違すると判断された場合、映像又は音声のうちいずれが相違するかをさらに判別する（ステップS61）。

【0075】映像に関して相違する場合には、映像照合リスト中に当該映像は一致しない旨を記録した後（ステップS62）、ステップS59に復帰して、対応映像の照合処理をスキップする。

【0076】また、音声に関して相違する場合には、音声照合リスト中に当該音声は一致しない旨を記録した後（ステップS63）、ステップS59に復帰して、対応音声の照合処理をスキップする。

【0077】特徴データベース中のすべてのデータについて、検索対象となるAVコンテンツとの相違を検証した後、本処理ルーチン全体を終了する。

【0078】このようにして照合の誤差を徐々に小さくしていくことで、最終的に特徴が近いコンテンツを高速に探索することができる。ステップS58までの工程では、明らかに相違する訳でないコンテンツがただ1つに絞り込まれた場合には、ステップS59以降の工程を合えて実行する必要はない。

【0079】また、被検信号に歪が多いことが想定される場合には、ステップS53～S58の工程において、違う程度を徐々に小さくして何回も繰り返すことで、誤差の最も小さいコンテンツを詳細に探索することができ

る。

【0080】以上、検索性能の改善について述べてきたが、最後に本発明に係る情報検索の応用例について説明しておく。

【0081】（1）まず第1の応用例は、放送の監視である。放送番組のスポンサと放送事業者との間では、一般に、広告の挿入に応じてスポンサから放送事業者に相当額（スポンサ料）を支払うという事業モデルが構築されている。

【0082】しかしながら、約束通りに広告が挿入されない可能性があるので、スポンサは実際の放送を監視する必要がある。このような場合、あらかじめ広告の特徴量をデータベースに登録しておき、放送を受信し放送コンテンツから特徴抽出してデータベースと照合する。

【0083】（2）さらに、不当に放送コンテンツが複製され放送に利用されていないかを、番組を制作した著作権者が監視するというアプリケーションを挙げることができる。

【0084】テープなどの記録メディアで複製又はその一部を持っている場合、アーカイブの中からその本編を探し出すという目的に、本発明に係る情報検索手法を使用することができる。

【0085】番組制作のための撮影時に、特徴抽出を同時に行い、特徴量のデータベースを作っておくことも賢明な方法である。この場合、カメラや撮影者の識別コード、日時、GPS（Global Positioning System）などを用いた緯度・経度などのデータも対応させて登録しておけば、さらに便利である。

【0086】（3）家庭においても、VTR（Video Tape Recorder）やディスク・レコーダで記録されたコンテンツから特徴抽出を行い、視聴履歴から関連する番組をお勧め版としてユーザに通知したり、自動録画したりするサービスも考えられる。特徴抽出は、複数のVTR、DVD（Digital Versatile Disc）プレーヤやその他の記録再生機器、チューナなどすべてのコンテンツが集まる受信機間で行うことで、合理化を図ることができる。

【0087】ユーザの視聴履歴から関連する番組をお勧め版としてユーザに教えるようなサービスでは、視聴者別のサービスなので、視聴者を特定するために受信機器の識別コードが必要となる。

【0088】さらに、家庭内の個人を特定するには、リモート・コマンドを個人毎に占有し、その識別コードを使用する。これらの識別コードも、サービス供給者に送る必要がある。受信時には、時刻やチャンネルが判ると検索が容易になる。また、信号の品質（受信時、再生時、再生のモード（例えば標準と3倍））が判ると、これをデータベース側に送り、照合時の許容度を適応的に制御することによって、検索速度又は検索の確実度を向上させることができる。

【0089】[追補]以上、特定の実施例を参照しながら、本発明について詳解してきた。しかしながら、本発明の要旨を逸脱しない範囲で当業者が該実施例の修正や代用を成し得ることは自明である。すなわち、例示という形態で本発明を開示してきたのであり、限定的に解釈されるべきではない。本発明の要旨を判断するためには、冒頭に記載した特許請求の範囲の欄を参酌すべきである。

【0090】

【発明の効果】以上詳記したように、本発明によれば、映像との音楽の各データからなるAVコンテンツを高速且つ確実に検索することができる、優れた情報検索装置及び方を提供することができる。

【0091】本発明に係る情報検索装置及び方法によれば、映像だけ、又は音声だけでAVコンテンツを検索するよりも迅速に検索することができる。また、系における歪を許容して検索する場合においても、より確実にAVコンテンツを検索することができる。

【図面の簡単な説明】

【図1】本発明の一実施形態に係る情報検索装置1の構成を模式的に示した図である。

【図2】情報検索装置1上で実行されるAVコンテンツの検索処理についての概略的な手順を示したフローチャートである。

【図3】情報検索装置1上で実行されるAVコンテンツ

の検索処理の具体的な手順の一例を示したフローチャートである。

【図4】情報検索装置1上で実行されるAVコンテンツの検索処理の具体的な手順の他の例を示したフローチャートである。

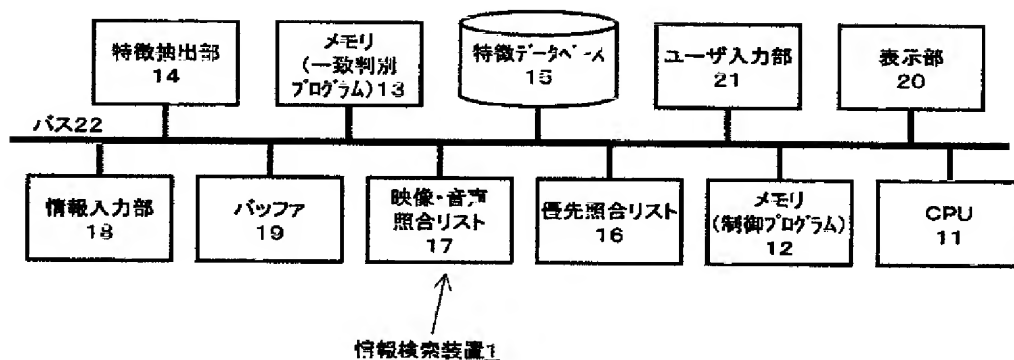
【図5】情報検索装置1上で実行されるAVコンテンツの検索処理の具体的な手順の他の例を示したフローチャートである。

【図6】ISRCのような楽曲のユニーク・コードを利用した音楽検索のメカニズムを示した図(従来例)である。

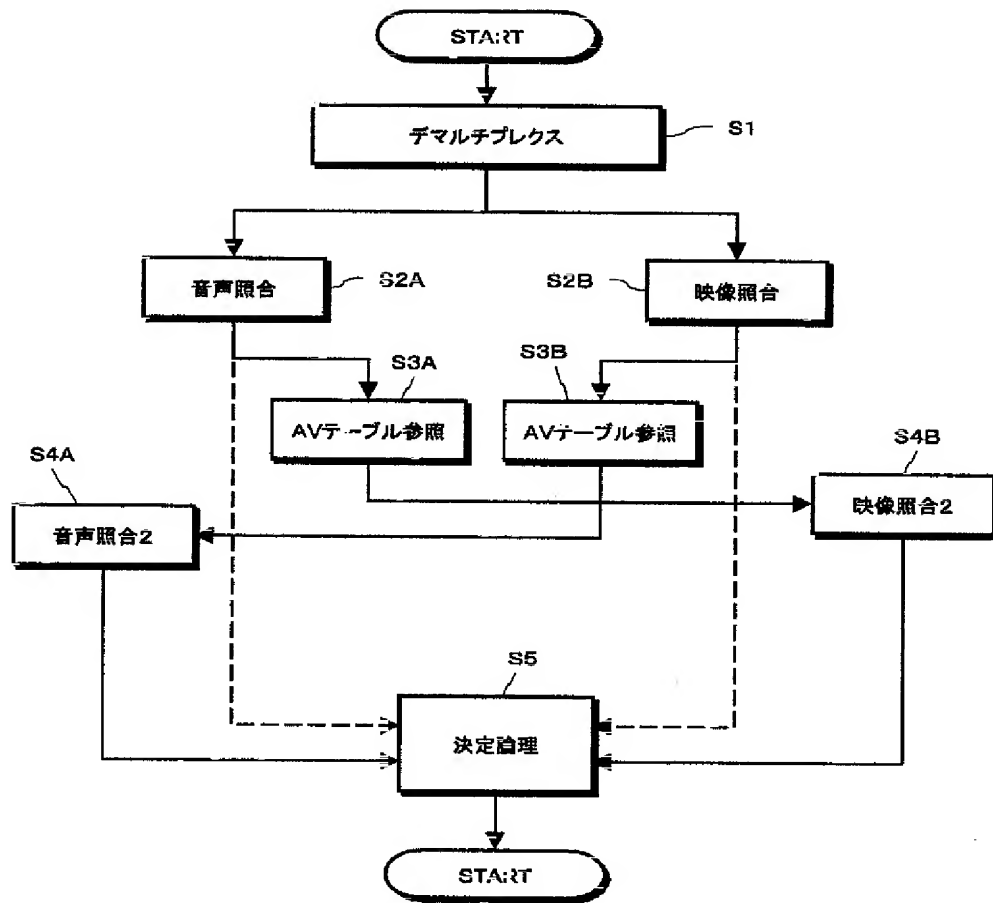
【符号の説明】

- 1…情報検索装置
- 11…CPU
- 12…メモリ(制御プログラム)
- 13…メモリ(一致判別プログラム)
- 14…特徴抽出部
- 15…特徴データベース
- 16…優先照合リスト
- 17…映像・音声照合リスト
- 18…情報入力部
- 19…バッファ
- 20…表示部
- 21…ユーザ入力部

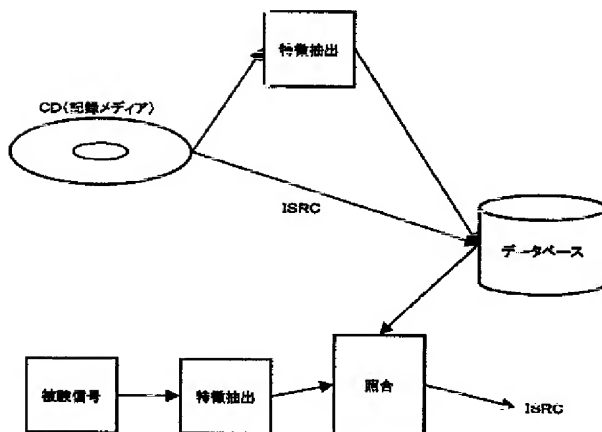
【図1】



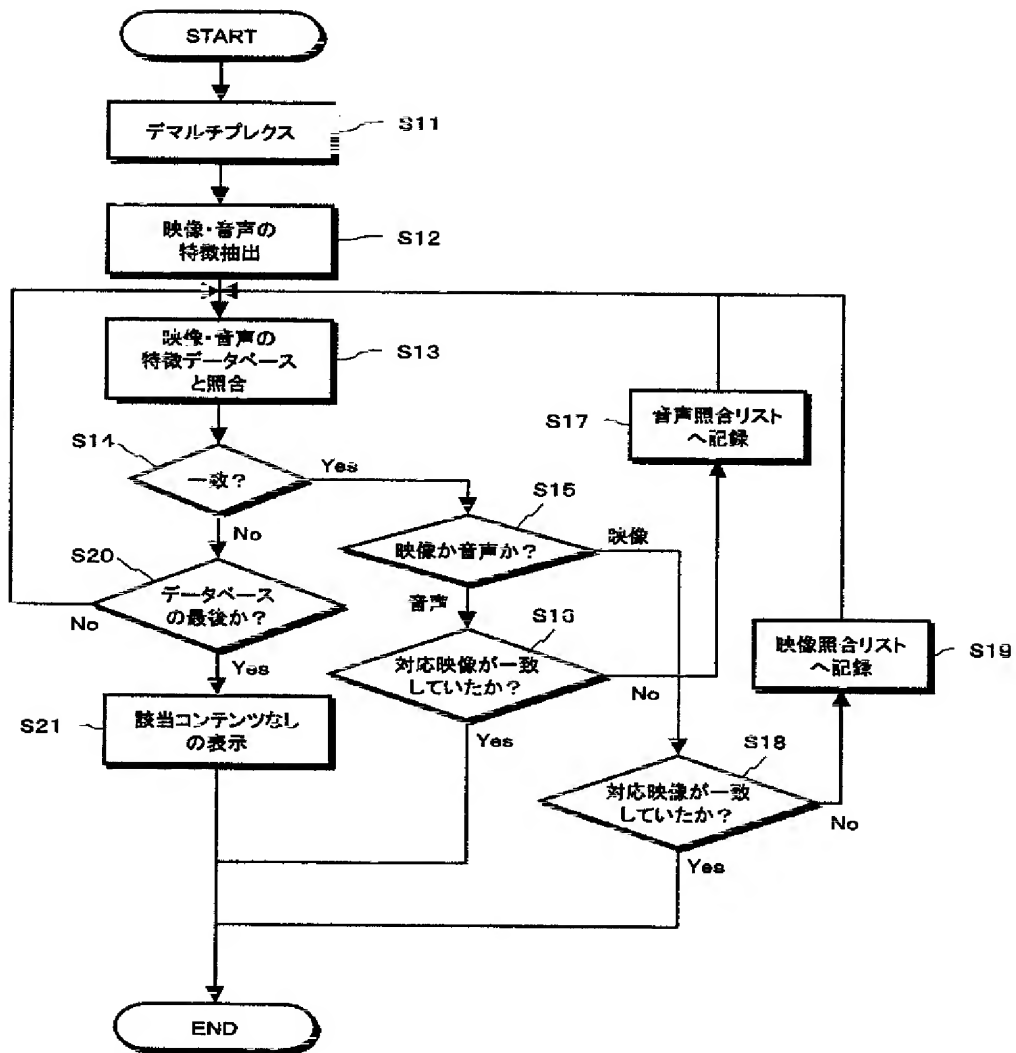
【図2】



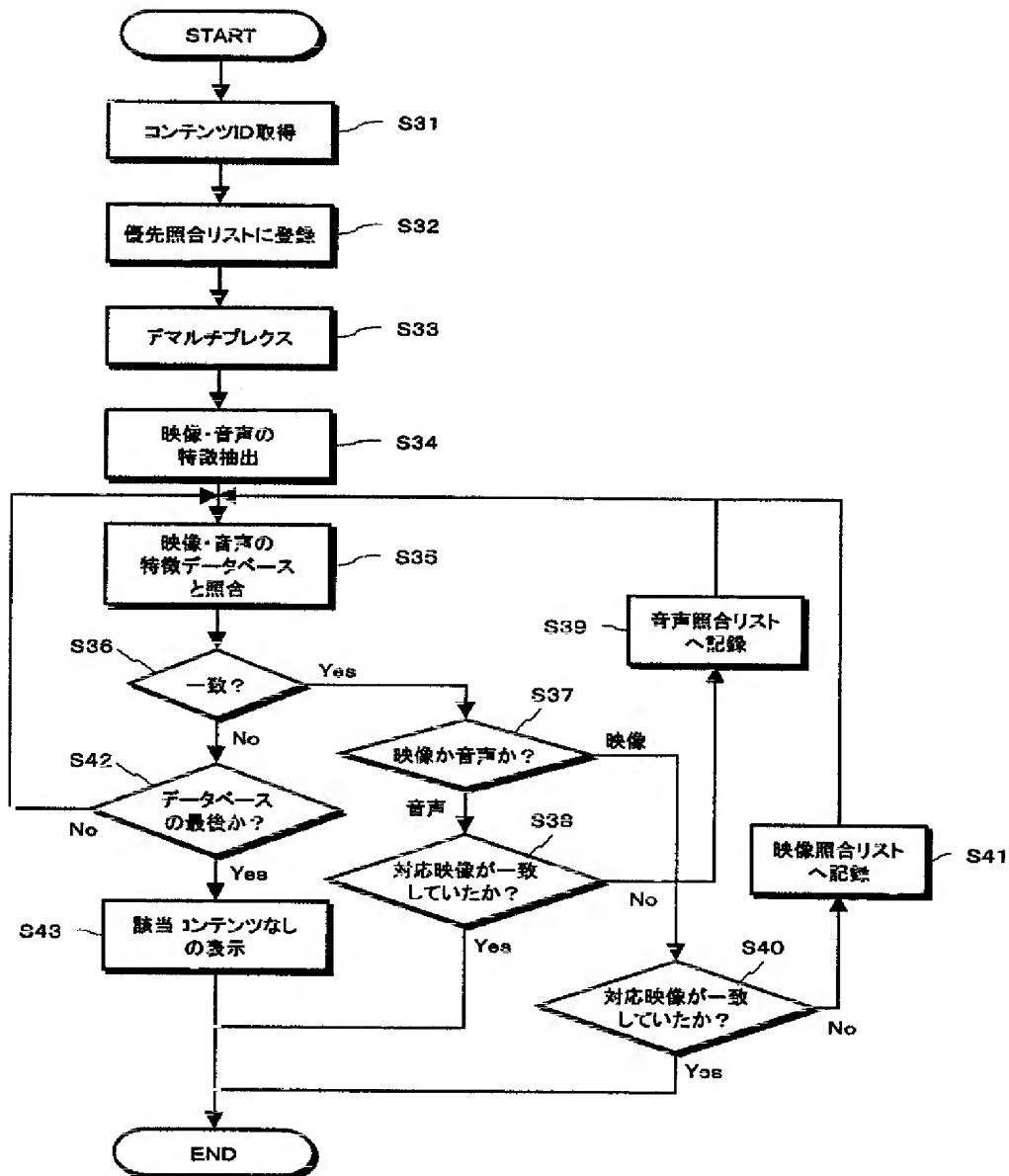
【図6】



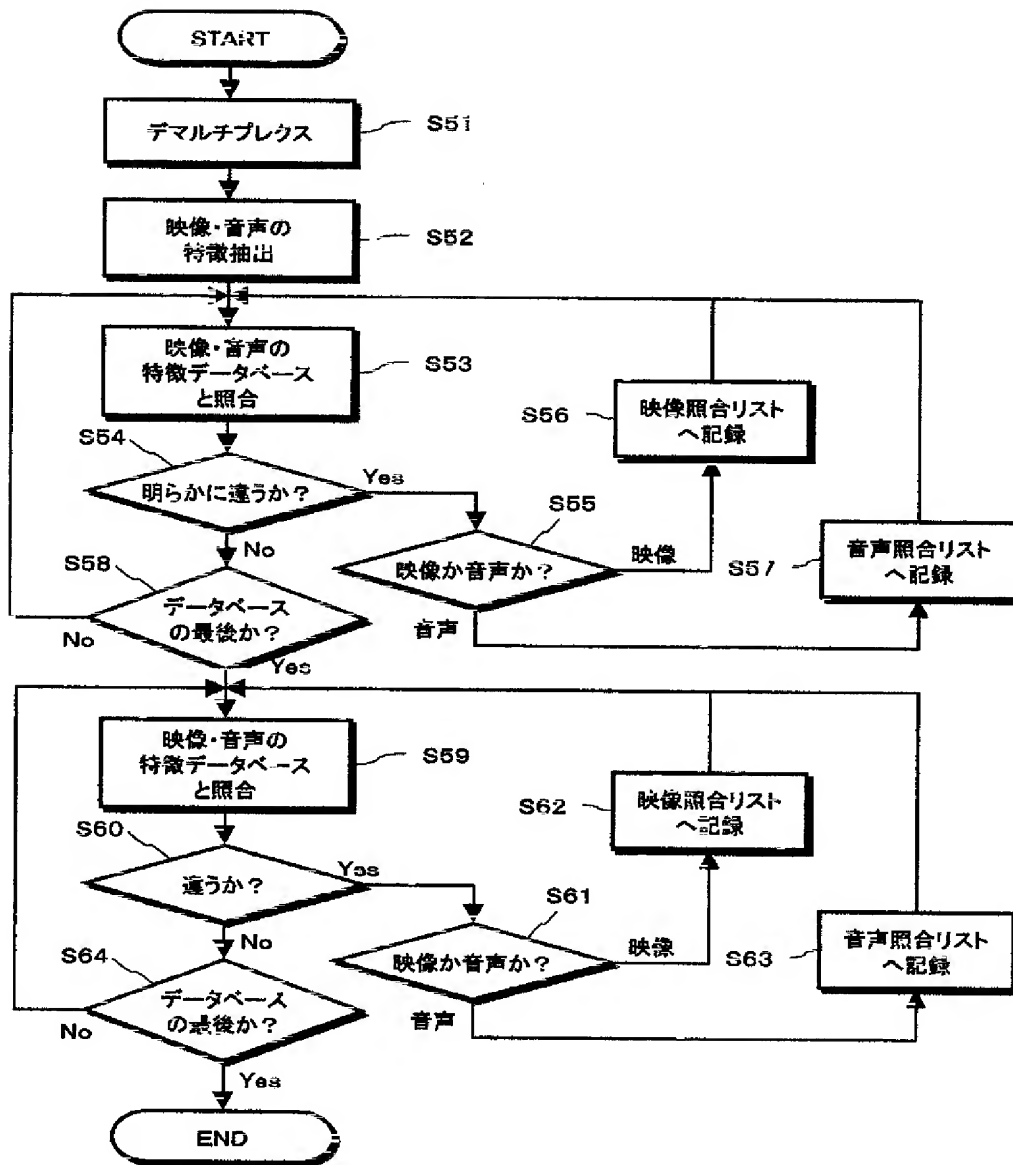
【図3】



【図4】



【図5】



フロントページの続き

(51)Int.Cl.<sup>7</sup>

G10L 15/28

H04N 5/76

識別記号

FI

G10L 3/00

551P

(参考)